

Supplementary Figure 1. Direct overlaps between genetic and physical interactions, while statistically significant, are limited in systematic data and probably biased.

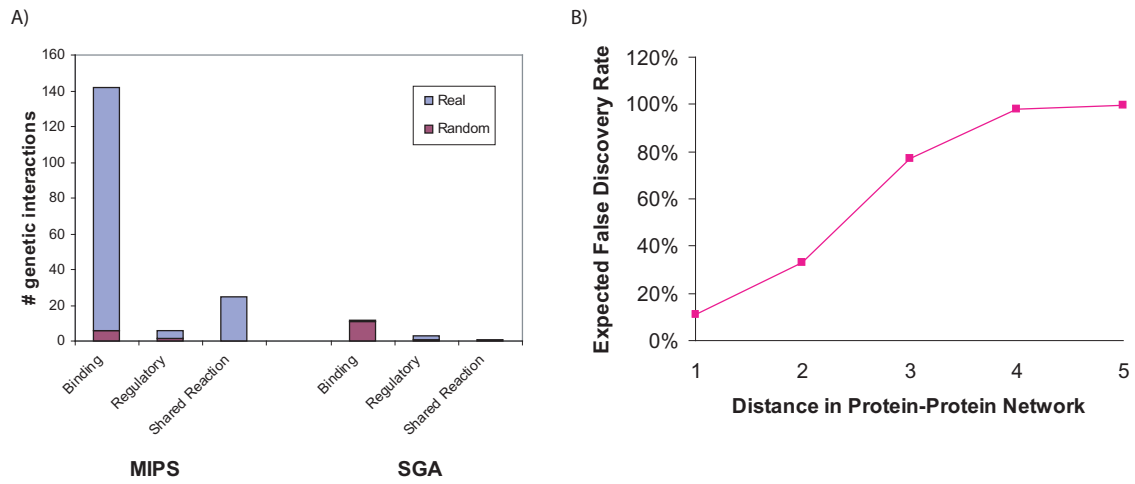
As a preliminary assessment of whether synthetic-lethal genetic interactions could be explained by physical interactions, we investigated whether proteins connected by genetic interactions were also at close proximity in the physical network. As shown in Figure S1 [a], genetic interactions were co-incident with a total of 189 physical interactions (154 protein binding, 9 regulatory, 26 metabolic). These counts were significant in comparison to randomized genetic networks (yielding 19.2 ± 6.9 overlapping physical interactions; mean \pm stdev). In the figure, results are tabulated separately for two types of genetic interactions (MIPS, SGA) and three types of physical network (protein-protein binding, protein-DNA regulatory, and shared-reaction metabolic).

However, further investigation suggested that much of this overlap might be due to bias in determination of the physical or genetic network. First, 93% (176/189) of the coincident genetic interactions were derived from small-scale studies curated by MIPS. This percentages was highly enriched ($p < 4 \times 10^{-65}$) compared to the relatively small percentage (26%) of genetic data measured in small-scale studies overall. Similarly, 87% (164/189) of the coincident physical interactions were identified in small-scale studies (as recorded by the DIP database).

Thus, the coincident physical interactions are biased towards small-scale studies, probably because physical interactions are sometimes tested explicitly as a follow-up to observing a genetic interaction. Direct correspondence between systematic genetic and physical interactions is much weaker (e.g., between SGA and protein-binding interactions in DIP)

Such conclusions hold even after extending the analysis from direct interactions to longer paths. As shown in Figure S1 [b], for each pair of synthetic-lethal proteins we recorded the length of the shortest path connecting these proteins in the protein-protein network. The False Discovery Rate of genetic/physical overlap is shown for genetic interactions connected by direct (length 1) or longer paths of protein-protein interactions (lengths 2-6). False Discovery Rate (FDR) is the expected percentage of these relationships that are spurious based on randomized networks (described further in the Methods). Genetic interactions in MIPS match a greater number of short paths than would be expected at random, while the number explained by physical paths of length > 3 (SGA: all lengths) is essentially no different than for random networks.

Thus, the number of synthetic-lethal pairs connected by paths of up to three protein-protein interactions was larger than expected. However, this trend was too weak to be used in identifying the physical cause of any particular synthetic-lethal interaction: Even limiting consideration to paths of length two, nearly a third of the paths were likely due to random chance.



Supplementary Figure 1: Direct overlaps between genetic and physical interactions, while statistically significant, are limited in systematic data and probably biased.